

11-64-CR

*TWO ALTERNATE PROOFS OF WANG'S LUNE FORMULA FOR SPARSE  
DISTRIBUTED MEMORY AND AN INTEGRAL APPROXIMATION*

Louis A. Jaeckel

Research Institute for Advanced Computer Science  
NASA Ames Research Center

RIACS TR 88.5

February 3, 1988

(NASA-CR-184555) TWO ALTERNATE PROOFS OF  
WANG'S LUNE FORMULA FOR SPARSE DISTRIBUTED  
MEMORY AND AN INTEGRAL APPROXIMATION (NASA)  
CSCL 12A  
35 P

N89-13214

Unclas  
G3/64 0175199

**RIACS**

Research Institute for Advanced Computer Science



**TWO ALTERNATE PROOFS OF WANG'S LUNE FORMULA FOR SPARSE  
DISTRIBUTED MEMORY AND AN INTEGRAL APPROXIMATION**

Louis A. Jaeckel

Research Institute for Advanced Computer Science  
NASA Ames Research Center

RIACS Technical Report 88.5

February 3, 1988

In Kanerva's Sparse Distributed Memory, writing to and reading from the memory are done in relation to spheres in an  $n$ -dimensional binary vector space. Thus it is important to know how many points are in the intersection of two spheres in this space. This paper gives two proofs of Wang's formula for spheres of unequal radii, and an integral approximation for the intersection in this case.

---

Work reported herein was supported in part by Cooperative Agreement NCC 2-408 between the National Aeronautics and Space Administration (NASA) and the Universities Space Research Association (USRA).



# **TWO ALTERNATE PROOFS OF WANG'S LUNE FORMULA FOR SPARSE DISTRIBUTED MEMORY AND AN INTEGRAL APPROXIMATION**

Louis A. Jaeckel

Research Institute for Advanced Computer Science

February 3, 1988

## **INTRODUCTION**

A. Wang of RIACS has derived a formula for the volume of a lune formed by two spheres of unequal radii in a binary vector space.<sup>1</sup> This paper gives two proofs of Wang's formula, and an integral approximation for the intersection in this case.

Let  $S$  be the set of all  $n$ -bit binary words. This set may be described as the set of all  $n$ -dimensional vectors, each of whose components is 0 or 1, and thus may be thought of geometrically as the set of all vertices of an  $n$ -dimensional unit cube embedded in an  $n$ -dimensional vector space, or as an  $n$ -dimensional vector space over the field consisting of the two elements 0 and 1. Any of the  $2^n$  points in  $S$  is like any other in the sense that there is an isometry of  $S$  which maps one point onto the other.

For two points  $x = (x_1, \dots, x_n)$  and  $y = (y_1, \dots, y_n)$  in  $S$ ,

---

<sup>1</sup>Wang's work will appear in a forthcoming RIACS technical report.

we let

$$d(x,y) = \sum_{i=1}^n |x_i - y_i| .$$

This is the Hamming distance, or L1 distance, between  $x$  and  $y$ . It is the number of coordinates for which  $x_i \neq y_i$ . This measure of distance is equivalent to the Euclidean, or L2, distance between points of  $S$ .

Kanerva (1984) uses this space as the address space for his Sparse Distributed Memory. Since writing to and reading from the memory are done in relation to spheres centered at the write or read address, we need to know how many points of  $S$  lie in the intersection of two such spheres. We will refer to the number of points in a subset of  $S$  as the "volume" of the subset. Kanerva (1984), p. 146, derived a formula for the volume of the lune formed by two spheres of equal radii, that is, the number of points that are in the first sphere but outside of the second. The volume of the intersection of the two spheres is then found by subtracting the volume of the lune from the volume of the first sphere. He also derived an integral approximation for the volumes of the lune and of the intersection, for spheres of equal radii. (Ibid., p.157)

Wang generalized Kanerva's formula for the volume of the lune to the case of spheres of unequal radii. This volume is of interest because in some applications of a Sparse Distributed Memory system we may want to use spheres of different radii when writing to and reading from the memory.

In this paper we give two alternate proofs of Wang's lune

formula. Each looks at the space  $S$  in a different way, and thus provides some insight into various aspects of its structure. We then derive an integral approximation for the volume of the intersection of spheres whose radii may be unequal, and show that for the case of equal radii, the approximation is equivalent to Kanerva's.

#### PROOF BY MOVING THE SPHERES AWAY FROM EACH OTHER

The proof in this section is based on the geometry of spheres in the space  $S$ , and is similar in spirit to the proofs of Kanerva (1984) and Wang. We will find the net increase in the lune as we move the centers of the spheres away from each other one unit at a time. The volume of the lune will then be the sum of these net increases.

We begin with a sphere of radius  $s$  which will remain stationary, and a sphere of radius  $r$  which will move. We will consider the lune consisting of those points which are in the  $r$ -sphere but not in the  $s$ -sphere. Suppose the distance between the centers of the spheres is  $k$ . If we move the center of the  $r$ -sphere one unit farther away from the center of the  $s$ -sphere, we will see that there is a net decrease in the number of points in the intersection of the two spheres, and an equal increase in the volume of the lune. Both Kanerva and Wang pointed out that, contrary to our usual intuition, when the  $r$ -sphere is moved in this way, some points may move from the lune back into the intersection; moreover, every second one-unit move results in no

net change.

We may assume without loss of generality that the center of the  $s$ -sphere is  $(0,0,\dots,0)$  and that the current center of the  $r$ -sphere is

$$x = (1,1,\dots,1,0,0,\dots,0) .$$

In this vector, the values for the first  $k$  coordinates are 1's, so that the distance between the centers is  $k$ . We will move the center of the  $r$ -sphere one unit by changing its  $(k+1)^{\text{st}}$  coordinate from 0 to 1, so that the new center will be

$$x' = (1,1,\dots,1,1,0,\dots,0) ,$$

a vector with  $k+1$  1's, and the new distance between the centers will be  $k+1$ .

We consider first the change in the  $r$ -sphere caused by moving its center. Many points in the sphere before it is moved remain in the sphere, while other points drop out and are replaced by an equal number of new points. In fact, we will see that there is a one-to-one correspondence between the points that drop out and their replacements. Note that we are not mapping each individual point of the sphere centered at  $x$  onto a corresponding point of the sphere centered at  $x'$ . Instead, the points in  $S$  will remain where they are, and we will think of the sphere as changing position. We can think of the points in  $S$  as being like the stationary light bulbs in the moving news bulletins on the sides of buildings, and the sphere as a message that appears to move as individual bulbs go on or off. Since we are interested in the net change in the intersection as the  $r$ -sphere moves, and the points which are in the sphere in both



its old and its new positions do not cause any change in the intersection, we will just ignore them and focus our attention on the points that drop out of the sphere and are replaced by new points.

For any point in  $S$ , if its  $k+1^{\text{st}}$  coordinate is 0, its distance from  $x'$  is one more than its distance from  $x$ . On the other hand, if its  $k+1^{\text{st}}$  coordinate is 1, its distance from  $x'$  is one less than its distance from  $x$ . Therefore, a point is in the  $r$ -sphere in both the old and the new positions of the sphere either if its distance from  $x$  is less than  $r$ , or if its distance from  $x$  is equal to  $r$  and its  $k+1^{\text{st}}$  coordinate is 1. These points do not produce any change in the lune or the intersection.

A point drops out of, or leaves, the  $r$ -sphere (when its center moves from  $x$  to  $x'$ ) if its distance from  $x$  is  $r$  and its  $k+1^{\text{st}}$  coordinate is 0. If we look at all of the coordinates of such a point, except the  $k+1^{\text{st}}$ , we see that the values for the coordinates disagree with  $x$  for exactly  $r$  of these  $n-1$  coordinates. Moreover, any assignment of 0's and 1's to these  $n-1$  coordinates so that exactly  $r$  of them disagree with  $x$ , together with assigning 0 to the  $k+1^{\text{st}}$  coordinate, will define such a leaving point.

Conversely, a point enters the  $r$ -sphere as the sphere moves if its distance from  $x'$  is  $r$  and its  $k+1^{\text{st}}$  coordinate is 1. As with the leaving points, the values for the coordinates of an entering point, for all of the coordinates except the  $k+1^{\text{st}}$ , disagree with  $x'$ , and therefore with  $x$ , on exactly  $r$  of these

$n-1$  coordinates. Also, any assignment of 0's and 1's to these  $n-1$  coordinates so that  $r$  of them disagree with  $x$ , together with assigning 1 to the  $k+1^{\text{st}}$  coordinate, will define an entering point.

Thus there is a one-to-one correspondence between the leaving points and the entering points: For any leaving point, change its  $k+1^{\text{st}}$  coordinate from 0 to 1 and leave the other coordinates the same, and we have an entering point, which we will sometimes call the "replacing point". The effect of moving the center of the  $r$ -sphere from  $x$  to  $x'$  is therefore to remove the leaving points from the sphere and to replace each of them with its corresponding replacing point.

Now we consider the effect of this operation on the intersection of the  $r$ -sphere and the  $s$ -sphere. The points that are in the  $r$ -sphere in both of its positions of course have no effect on the intersection. Moreover, if a leaving point and its replacing point are both in the  $s$ -sphere, they have no net effect on the number of points in the intersection. The same is true if neither of them is in the  $s$ -sphere. The points that do have an effect are those leaving points which are in the  $s$ -sphere but whose replacing points are not. We note that it is impossible for a leaving point to be outside the  $s$ -sphere and its replacing point to be in it, because the distance of a replacing point from the center of the  $s$ -sphere is always one greater than the distance of the corresponding leaving point from the center of the  $s$ -sphere.

If the leaving point is in the  $s$ -sphere and its replacing

point is not, then the distance of the leaving point from the center of the  $s$ -sphere must be exactly  $s$ , so that the distance of the replacing point from the center will be  $s+1$ . The decrease in the volume of the intersection -- and the corresponding increase in the volume of the lune -- is therefore the number of leaving-replacing point pairs satisfying this condition.

Let  $a$  be the number of 0's among the first  $k$  coordinates of a given leaving point, and let  $b$  be the number of 1's among the  $k+2^{\text{nd}}$  to  $n^{\text{th}}$  coordinates. There are  $n-1-k$  coordinates in the latter set. The  $k+1^{\text{st}}$  coordinate of the point is 0. Since the point is a leaving point, its distance from  $x$  is  $r$ . Therefore, since  $x$  consists of  $k$  1's followed by  $n-k$  0's,

$$a + b = r .$$

If this leaving point is in the  $s$ -sphere but its replacing point is not, then its distance from the center of the  $s$ -sphere is  $s$ , so

$$(k - a) + b = s .$$

Solving these equations for  $a$  and  $b$ , we find:

$$a = \frac{k+r-s}{2}$$

and

$$b = \frac{r+s-k}{2} .$$

If  $k$  is congruent to  $s-r \pmod{2}$ , then  $a$  and  $b$  are integers, and the number of ways of assigning  $a$  0's among the first  $k$  coordinates and  $b$  1's among the last  $n-1-k$  coordinates is

$$\left( \frac{k}{\frac{k+r-s}{2}} \right) \cdot \left( \frac{n-1-k}{\frac{r+s-k}{2}} \right).$$

This is the number of leaving points satisfying the above conditions. Therefore this is the increase in the number of points in the lune when the distance between the centers is increased from  $k$  to  $k+1$ . On the other hand, if  $k$  is not congruent to  $s-r \pmod{2}$ , then  $a$  and  $b$  are not integers. In that case no points satisfy the above conditions, and there is no net change in the volume of the lune.

Consider the lune consisting of those points in the  $r$ -sphere but outside of the  $s$ -sphere, and assume that  $r \leq s$ . We begin with the center of the  $r$ -sphere at the center of the  $s$ -sphere, and move the former away from the latter one unit at a time. At first, the  $r$ -sphere is completely contained in the  $s$ -sphere, so the volume of the lune is 0. This continues to be true as long as  $k < s-r$ . When  $k = s-r$ , at which point we increase the distance between the centers from  $k$  to  $k+1 = s-r+1$ , we finally have some points in the lune.

To give a formula for the volume of the lune when the distance between the centers is  $d$ , we let  $k$  increase from  $s-r$  to  $d-1$  and add the net increases at each step. Note that the term in the sum for  $k = d-1$  corresponds to increasing the distance between the centers from  $d-1$  to  $d$ . Since for every second value of  $k$  the volume does not change, these values of  $k$  are omitted. The volume of the lune for  $r \leq s$  is therefore

$$L = \sum_{\substack{k = s-r \\ k \equiv s-r \\ (\text{mod } 2)}}^{d-1} \binom{k}{\frac{k+r-s}{2}} \cdot \binom{n-1-k}{\frac{r+s-k}{2}} .$$

Note that by the symmetry property discussed in the next section, the first term in the summand could be written as

$$\binom{k}{\frac{k+s-r}{2}} .$$

The volume of the intersection is then the volume of the  $r$ -sphere, which is

$$\sum_{j=0}^r \binom{n}{j} ,$$

minus the volume of the lune.

If  $r = s$ , this lune formula is the same as the one derived by Kanerva (1984), p. 146. If  $r > s$ , the individual terms in the sum above correctly give the increase in the lune as the spheres move apart, but to find the volume of the lune, we must sum these terms over  $k = 0$  to  $d-1$ , and add to that the volume of the lune when the spheres have the same center, which in this case is

$$\sum_{j=s+1}^r \binom{n}{j} .$$

If we continue to move the  $r$ -sphere until it becomes disjoint from the  $s$ -sphere, we obtain a formula analogous to Corollary 1a of Kanerva (1984), p. 152. In the lune formula above, assume that  $r+s+1 \leq n$ , and let  $d = r+s+1$ , which is the smallest distance between the centers for which the spheres are disjoint. Then the lune is the entire  $r$ -sphere, and we have the following generalization of Kanerva's corollary:

$$L = \sum_{\substack{k = s-r \\ k \equiv s-r \\ (\text{mod } 2)}}^{r+s} \left( \frac{k}{\frac{k+r-s}{2}} \right) \cdot \left( \frac{n-1-k}{\frac{r+s-k}{2}} \right) = \sum_{j=0}^r \binom{n}{j}.$$

These formulas are related to some combinatorial results for random walks which may be found in Feller (1957), p. 65-87.

### PROOF BY INDUCTION ON THE DIMENSION OF S

We will now look at S from a different perspective, and derive a somewhat more general formula for the volume of the lune, which is valid for any  $r$  and  $s$ . The proof is by induction on  $n$ , the dimension of S.

Let  $L(n,d,r,s)$  be the volume of a lune in an  $n$ -dimensional space S, that is, the number of points in a sphere of radius  $r$  but outside a sphere of radius  $s$ , where  $d$  is the distance between the centers of the spheres. We will show that for all  $r$  and  $s$  and all  $0 \leq d \leq n$ ,  $L(n,d,r,s)$  is equal to the following sum:

$$\Sigma(n,d,r,s) = \sum_{\substack{k = s-r \\ k \equiv s-r \\ (\text{mod } 2)}}^{d-1} \left( \frac{k}{\frac{k+r-s}{2}} \right) \cdot \left( \frac{n-1-k}{\frac{r+s-k}{2}} \right).$$

If  $r \leq s$ , this sum is the same as in the previous section. However, the formula is also valid when  $r > s$ , in which case  $k$  begins with negative values. (Since binomial coefficients with negative  $k$  do not have the symmetry property given below, we must write the first term in the summand as it appears above, rather than in the alternate form given in the last section.)

Binomial coefficients may be defined for any value for the upper parameter, positive, negative, or zero, and for any integer value for the lower parameter. Since such values may now appear in the formula above, we define binomial coefficients as follows:

$$\binom{a}{b} = \frac{a(a-1)\cdots(a-b+1)}{b!} \quad \text{for } b > 0$$

$$\binom{a}{0} = 1$$

$$\binom{a}{b} = 0 \quad \text{for } b < 0.$$

As usual,  $0! = 1$ . For any  $a$  and  $b$ , the fundamental recursive property of binomial coefficients holds:

$$\binom{a}{b} = \binom{a-1}{b-1} + \binom{a-1}{b}.$$

For  $a \geq 0$ , we have the symmetry property

$$\binom{a}{b} = \binom{a}{a-b},$$

an example of which appeared in the last section. It follows from this property that if  $b > a \geq 0$ , then  $\binom{a}{b} = 0$ . Finally, if  $a < 0$ , we can see from the definition above that

$$\binom{a}{b} = (-1)^b \binom{-a+b-1}{b}.$$

These definitions and properties are given in Feller (1957), p. 48.

We will derive two recursive formulas for  $L$ , which will allow us to begin with any  $n$  and  $d$  such that  $0 \leq d \leq n$ , and reduce the dimension of the space down to  $n = 0$ . (A zero-dimensional vector space consists of just one point.)

Let  $c_r$  and  $c_s$  be the centers of the  $r$ -sphere and of the  $s$ -sphere, respectively. The lune is then the set of points whose distance from  $c_r$  is less than or equal to  $r$ , and whose

distance from  $c_s$  is greater than  $s$ .

Suppose first that  $n > d \geq 0$ . Then, on at least one coordinate, say on the  $n^{\text{th}}$ , the components of the two center vectors agree. Assume their common value on the  $n^{\text{th}}$  coordinate is 0. We divide the space into two halves: Let  $S_0$  be those points whose  $n^{\text{th}}$  coordinate is 0, and let  $S_1$  be those points for which it is 1. Each half may be viewed as an  $(n-1)$ -dimensional vector space, simply by dropping the  $n^{\text{th}}$  coordinate from each point. The distance between any two points within each half, computed by summing over the first  $n-1$  coordinates, is the same as it is in  $S$ . Since both  $c_r$  and  $c_s$  are in  $S_0$ , it follows that the points in  $S_0$  which are in the lune in  $S$  form a lune in  $S_0$ , and the number of such points is  $L(n-1, d, r, s)$ .

Let  $c'_r$  and  $c'_s$  be the points in  $S_1$  whose first  $n-1$  coordinates are the same as those of  $c_r$  and  $c_s$ . The distance from any point in  $S_1$  to either of these points is one less than the distance from the point to  $c_r$  or to  $c_s$ . Therefore, a point in  $S_1$  is in the  $r$ -sphere iff its distance from  $c'_r$  is less than or equal to  $r-1$ , and it is outside of the  $s$ -sphere iff its distance from  $c'_s$  is greater than  $s-1$ . Since the distance between  $c'_r$  and  $c'_s$  is  $d$ , it follows that the points in  $S_1$  that are in the lune in  $S$  form a lune in  $S_1$ , and the number of such points is  $L(n-1, d, r-1, s-1)$ .

Therefore,

$$L(n, d, r, s) = L(n-1, d, r, s) + L(n-1, d, r-1, s-1) .$$

Now suppose that  $n \geq d > 0$ . Then, on at least one coordinate, say on the  $n^{\text{th}}$ , the components of the two center



vectors differ. Assume that the  $n^{\text{th}}$  coordinate of  $c_r$  is 0, and that for  $c_s$  it is 1. We divide  $S$  into two halves as before.

We consider  $S_0$  first. This set contains  $c_r$  but not  $c_s$ . Let  $c'_s$  be the point in  $S_0$  whose first  $n-1$  coordinates are the same as those of  $c_s$ . Since the distance from any point in  $S_0$  to  $c'_s$  is one less than its distance to  $c_s$ , a point in  $S_0$  is in the lune in  $S$  iff its distance from  $c_r$  is less than or equal to  $r$ , and its distance from  $c'_s$  is greater than  $s-1$ . These points form a lune in  $S_0$ , and since the distance between  $c_r$  and  $c'_s$  is  $d-1$ , the number of such points is  $L(n-1, d-1, r, s-1)$ .

By a similar argument, the points in  $S_1$  which are in the lune in  $S$  form a lune in  $S_1$ , and the number of such points is  $L(n-1, d-1, r-1, s)$ . Therefore,

$$L(n, d, r, s) = L(n-1, d-1, r, s-1) + L(n-1, d-1, r-1, s) .$$

Using these two recursive formulas, we can reduce the problem down to a zero-dimensional space, preserving the inequalities  $0 \leq d \leq n$  at each step. Or, conversely, we can begin with  $n = d = 0$  and build up to any  $n$  and  $d$  inductively. We will give the boundary conditions later. It is important to note that at some stages of this process, the values of  $r$  and  $s$  may be greater than  $n$ , or even less than 0. Although spheres with such radii may not have much geometric meaning, they are still well-defined algebraically, since a sphere is defined as the set of points whose coordinates satisfy a certain inequality. This is true even for  $n = 0$ , in which case the single point in the space has no coordinates, since a

sum with no terms in it is defined to be zero.

To prove that the sum  $\Sigma(n,d,r,s)$  given above is the volume of the lune, we must show that it satisfies the two recursive formulas, and that it also satisfies the appropriate boundary conditions for  $n = 0$ . The recursive formulas and the boundary conditions together uniquely determine all of the values of any function satisfying them. We begin by showing that  $\Sigma$  satisfies the two recursive formulas for all  $r$  and  $s$ , and for all  $n$  and  $d$  satisfying the inequalities below.

For the first formula, assume that  $n > d \geq 0$ . By the binomial coefficient formula above,

$$\begin{aligned} \binom{\frac{n-1-k}{2}}{\frac{r+s-k}{2}} &= \binom{\frac{n-2-k}{2}}{\frac{r+s-k}{2}} + \binom{\frac{n-2-k}{2}}{\frac{r+s-k}{2} - 1} \\ &= \binom{\frac{(n-1)-1-k}{2}}{\frac{r+s-k}{2}} + \binom{\frac{(n-1)-1-k}{2}}{\frac{(r-1)+(s-1)-k}{2}}. \end{aligned}$$

Therefore, inserting these terms into the sum and summing over  $k \equiv s-r \pmod{2}$ , we have

$$\begin{aligned} \Sigma(n,d,r,s) &= \sum_{\substack{k = s-r \\ k \equiv s-r \\ (\text{mod } 2)}}^{d-1} \binom{\frac{k}{2}}{\frac{k+r-s}{2}} \cdot \binom{\frac{n-1-k}{2}}{\frac{r+s-k}{2}} \\ &= \sum_{\substack{k = s-r \\ k \equiv s-r \\ (\text{mod } 2)}}^{d-1} \binom{\frac{k}{2}}{\frac{k+r-s}{2}} \cdot \binom{\frac{(n-1)-1-k}{2}}{\frac{r+s-k}{2}} \\ &\quad + \sum_{\substack{k = s-r \\ k \equiv s-r \\ (\text{mod } 2)}}^{d-1} \binom{\frac{k}{2}}{\frac{k+(r-1)-(s-1)}{2}} \cdot \binom{\frac{(n-1)-1-k}{2}}{\frac{(r-1)+(s-1)-k}{2}}. \end{aligned}$$

Since the initial value of  $k$  in the last sum above may be written as  $(s-1)-(r-1)$ , we have

$$\Sigma(n,d,r,s) = \Sigma(n-1,d,r,s) + \Sigma(n-1,d,r-1,s-1) .$$

For the second recursive formula, assume  $n \geq d > 0$ .

Applying the binomial coefficient formula again, and letting  $j = k-1$ , we have

$$\begin{aligned} \left( \frac{k}{\frac{k+r-s}{2}} \right) &= \left( \frac{k-1}{\frac{k+r-s}{2}} \right) + \left( \frac{k-1}{\frac{k+r-s}{2} - 1} \right) \\ &= \left( \frac{j+r-(s-1)}{2} \right) + \left( \frac{j+(r-1)-s}{2} \right) . \end{aligned}$$

The other binomial coefficient in the summand can be written as

$$\left( \frac{n-1-k}{\frac{r+s-k}{2}} \right) = \left( \frac{(n-1)-1-j}{\frac{r+(s-1)-j}{2}} \right) = \left( \frac{(n-1)-1-j}{\frac{(r-1)+s-j}{2}} \right) .$$

Putting these together and summing over  $k \equiv s-r \pmod{2}$ , which is equivalent to  $j \equiv (s-1)-r \equiv s-(r-1)$ , we have

$$\begin{aligned} \Sigma(n,d,r,s) &= \sum_{\substack{k = s-r \\ k \equiv s-r \\ (\text{mod } 2)}}^{d-1} \left( \frac{k}{\frac{k+r-s}{2}} \right) \cdot \left( \frac{n-1-k}{\frac{r+s-k}{2}} \right) \\ &= \sum_{\substack{j = (s-1)-r \\ j \equiv (s-1)-r \\ (\text{mod } 2)}}^{d-2} \left( \frac{j+r-(s-1)}{2} \right) \cdot \left( \frac{(n-1)-1-j}{\frac{r+(s-1)-j}{2}} \right) \\ &\quad + \sum_{\substack{j = s-r-1 \\ j \equiv s-(r-1) \\ (\text{mod } 2)}}^{d-2} \left( \frac{j+(r-1)-s}{2} \right) \cdot \left( \frac{(n-1)-1-j}{\frac{(r-1)+s-j}{2}} \right) . \end{aligned}$$

In the last sum above, if  $j = s-r-1$ , then the first term in the summand is 0, so we can begin the sum with the next available

value of  $j$ , which is two greater:  $j = s - (r-1)$ . Therefore,

$$\Sigma(n,d,r,s) = \Sigma(n-1,d-1,r,s-1) + \Sigma(n-1,d-1,r-1,s) .$$

We have now shown that the sum satisfies the two recursive formulas. To establish boundary conditions, we consider the case  $n = d = 0$ , with any values for  $r$  and  $s$ . In this case we have a zero-dimensional vector space containing one point, which is the center of both spheres. Since the  $r$ -sphere is the set of points whose distance from the center is less than or equal to  $r$ , the sphere contains one point if  $r \geq 0$ , and it is empty if  $r < 0$ . Similarly, the complement of the  $s$ -sphere is empty if  $s \geq 0$ , and it contains one point if  $s < 0$ . Therefore, the number of points in the lune is 0 or 1:

$$\begin{aligned} L(0,0,r,s) &= 1 \quad \text{if } r \geq 0 \text{ and } s < 0 \\ &= 0 \quad \text{otherwise .} \end{aligned}$$

We rewrite the sum for  $n = d = 0$  by letting  $i = \frac{k+r-s}{2}$ , so that  $k = 2i-r+s$ . The sum now becomes

$$\begin{aligned} \Sigma(0,0,r,s) &= \sum_{\substack{k = s-r \\ k \equiv s-r \\ (\text{mod } 2)}}^{-1} \left( \frac{k}{\frac{k+r-s}{2}} \right) \cdot \left( \frac{-1-k}{\frac{r+s-k}{2}} \right) \\ &= \sum_{i=0}^I \left( \frac{2i-r+s}{i} \right) \cdot \left( \frac{-1-2i+r-s}{r-i} \right) . \end{aligned}$$

This sum is over consecutive integers  $i$ , from 0 to  $I$ , which is the greatest integer in  $\frac{r-s-1}{2}$ . Since  $k = 2i-r+s$  is always negative in this sum, we can rewrite the first binomial coefficient as

$$\left( \frac{2i-r+s}{i} \right) = (-1)^i \left( \frac{r-s-i-1}{i} \right) .$$

Now let  $t = -s-1$ , and the sum becomes

$$A(r,t) = \sum_{i=0}^I (-1)^i \binom{r+t-i}{i} \cdot \binom{r+t-2i}{r-i},$$

where  $I = \left\lfloor \frac{r+t}{2} \right\rfloor$ .

Since  $s < 0$  is equivalent to  $t \geq 0$ , we must show that this sum is 1 if  $r \geq 0$  and  $t \geq 0$ , and is 0 otherwise. If  $r+t < 0$ , there are no terms in the sum, and the sum is 0. If  $r+t \geq 0$ , then for all terms in the sum,  $i \leq \frac{r+t}{2}$ , so

$$(r-i) + (t-i) = r + t - 2i \geq 0,$$

and by the symmetry property,

$$\binom{r+t-2i}{r-i} = \binom{r+t-2i}{t-i}.$$

If  $r < 0$ , then for all terms in the sum,  $r-i < 0$ , so this binomial coefficient is 0, and the sum is 0. Similarly, if  $t < 0$ , then the sum is 0. So the sum is 0 if either  $r$  or  $t$  is negative.

Assume from now on that  $r \geq 0$  and  $t \geq 0$ . We see from the symmetry property above that the summand is 0 if  $i > r$  or  $i > t$ . Therefore we can redefine  $I$ , the upper limit for  $i$ , to be  $I = \min(r,t)$ . This is less than or equal to the original value of  $I$ .

Suppose that either  $r = 0$  or  $t = 0$ . Then the sum has only one term, the term for which  $i = 0$ , and in either case that term is 1. Therefore,  $A(0,t) = A(r,0) = 1$ .

For any  $r \geq 0$  and  $t \geq 0$ , and any  $i \leq I = \min(r,t)$ :

$$\binom{r+t-i}{i} \cdot \binom{r+t-2i}{r-i} = \frac{(r+t-i)!}{i! (r+t-2i)!} \cdot \frac{(r+t-2i)!}{(r-i)! (t-i)!} \cdot \frac{r!}{r!}$$

$$\begin{aligned}
&= \frac{r!}{i! (r-i)!} \cdot \frac{(r+t-i)!}{r! (t-i)!} \\
&= \binom{r}{i} \cdot \binom{r+t-i}{r} .
\end{aligned}$$

The sum is now

$$A(r,t) = \sum_{i=0}^I (-1)^i \binom{r}{i} \cdot \binom{r+t-i}{r} .$$

We will show that this sum is 1 for all  $r \geq 0$  and  $t \geq 0$ .

The proof is by induction on the quantity  $M = r+t$ . We have already shown that  $A(0,t) = A(r,0) = 1$  for non-negative  $r$  and  $t$ . We therefore assume that  $A(r,t) = 1$  for all non-negative  $r$  and  $t$  such that  $r+t < M$ .

Let  $r > 0$  and  $t > 0$  be such that  $r+t = M$ . Applying the binomial coefficient formula twice to the summand, we have

$$\begin{aligned}
\binom{r}{i} \cdot \binom{r+t-i}{r} &= \binom{r}{i} \cdot \binom{r+t-i-1}{r} + \binom{r}{i} \cdot \binom{r+t-i-1}{r-1} \\
&= \binom{r}{i} \cdot \binom{r+t-i-1}{r} + \binom{r-1}{i} \cdot \binom{r+t-i-1}{r-1} + \binom{r-1}{i-1} \cdot \binom{r+t-i-1}{r-1} .
\end{aligned}$$

Multiplying each of these terms by  $(-1)^i$  and summing over  $i \geq 0$ , we find that the first of these three terms gives  $A(r,t-1)$  and the second term gives  $A(r-1,t)$ , assuming that the upper limit for  $i$  is correct. For the first sum, if  $r < t$ , the upper limit remains  $I = r$ ; if  $r \geq t$ , the summand is 0 for  $i = I = t$ , so the upper limit becomes  $t-1$ . For the second sum, if  $r > t$ , the upper limit remains  $I = t$ ; if  $r \leq t$ , the summand is 0 for  $i = I = r$ , so the upper limit becomes  $r-1$ . So for both sums the upper limit is the minimum of the two parameters.

For the third term in the expression above, we let  $j = i-1$ . Multiplying by  $(-1)^i$  and summing over  $i \geq 0$  gives

$$(-1) \sum_{j=-1}^{I-1} (-1)^j \binom{r-1}{j} \cdot \binom{r-1+t-1-j}{r-1}.$$

Since the summand is 0 for  $j = -1$ , we can change the initial value of  $j$  to 0. The upper limit for  $j$  is

$$I - 1 = \min(r, t) - 1 = \min(r-1, t-1).$$

Therefore this sum is  $-A(r-1, t-1)$ .

We now have

$$A(r, t) = A(r, t-1) + A(r-1, t) - A(r-1, t-1).$$

By induction on  $M = r+t$ , each of the three terms on the right is 1. Therefore,  $A(r, t) = 1+1-1 = 1$ , and the proof is complete.

As in the previous section, the volume of the intersection of the spheres is the volume of the  $r$ -sphere minus the volume of the lune.

A special case of the lune formula, as derived in this section, is the case where  $r > s$  and  $d = 0$ . The spheres are then concentric, and the lune is a kind of spherical shell, whose volume is

$$\sum_{j=s+1}^r \binom{n}{j}.$$

By the lune formula, this volume is equal to

$$\sum_{\substack{k = s-r \\ k \equiv s-r \\ (\text{mod } 2)}}^{-1} \binom{k}{\frac{k+r-s}{2}} \cdot \binom{n-1-k}{\frac{r+s-k}{2}}.$$

Unlike the sums derived in the previous section, this sum, in which all values of  $k$  are negative, is an alternating sum. To see this, we let  $i = \frac{k+r-s}{2}$ , as we did earlier in this section, and rewrite the sum as

$$\begin{aligned}
& \sum_{i=0}^I \binom{2i-r+s}{i} \cdot \binom{n-1+r-s-2i}{r-i} \\
&= \sum_{i=0}^I (-1)^i \binom{r-s-i-1}{i} \cdot \binom{n-1+r-s-2i}{r-i} = \sum_{j=s+1}^r \binom{n}{j},
\end{aligned}$$

where  $I = \left\lfloor \frac{r-s-1}{2} \right\rfloor$ .

Finally, if we let  $s = -1$  in the sum above, the  $s$ -sphere is empty, and we have an alternating sum for the volume of a sphere:

$$\sum_{i=0}^I (-1)^i \binom{r-i}{i} \cdot \binom{n+r-2i}{r-i} = \sum_{j=0}^r \binom{n}{j},$$

where  $I = \left\lfloor \frac{r}{2} \right\rfloor$ .

## AN INTEGRAL APPROXIMATION

Kanerva (1984) derived an integral approximation for the volume of the intersection of two spheres with the same radius. We will use a different method to derive an approximation for the case of two spheres whose radii may be unequal. The two approximations will be shown to be equivalent if the radii are equal.

Suppose we have a sphere of radius  $s$  with center  $c_s = (0, 0, \dots, 0)$ , and a sphere of radius  $r$  with center  $c_r = (1, 1, \dots, 1, 0, \dots, 0)$ , where the first  $d$  components of  $c_r$  are 1's, so that the distance between the centers is  $d$ . For any point in  $S$ , let  $y$  be the number of 0's among the first  $d$



coordinates, and let  $x$  be the number of 0's among the last  $n-d$  coordinates. The number of points in  $S$  having a given  $x$  and a given  $y$  is

$$\binom{n-d}{x} \cdot \binom{d}{y} .$$

If we choose a point in  $S$  at random, as if we were to toss a fair coin  $n$  times to determine the values for each of the coordinates, then  $x$  and  $y$  will be independent binomial random variables, where  $x$  has mean  $\frac{n-d}{2}$  and standard deviation  $\frac{\sqrt{n-d}}{2}$ , and  $y$  has mean  $\frac{d}{2}$  and standard deviation  $\frac{\sqrt{d}}{2}$ . The joint probability distribution of  $x$  and  $y$  is

$$\binom{n-d}{x} \cdot \binom{d}{y} \cdot 2^{-n} .$$

For moderately large  $d$  and  $n-d$ , this joint distribution may be closely approximated by a bivariate normal distribution.

For any point in  $S$ , its distance from  $c_s$  is the number of 1's among its coordinates:

$$(d - y) + (n - d - x) = n - x - y .$$

Its distance from  $c_r$  is the number of coordinates on which it differs from  $c_r$ :

$$y + (n - d - x) = n - d - x + y .$$

We see that both distances are functions of  $x$  and  $y$ . We can therefore group the points of  $S$  according to their values of  $x$  and  $y$ , and thus represent  $S$  by lattice points in the  $XY$  plane, as shown in Figure 1, where each point in the plane is given a weight according to the number of points of  $S$  it represents. This weight is  $2^n$  times the probability distribution above.

The  $s$ -sphere is the set of points satisfying

$$n - x - y \leq s ,$$

which may be written as

$$y \geq n - s - x .$$

This sphere is represented in Figure 1 as the region in the XY plane on and above the line  $y = n - s - x$ . The r-sphere is the set of points satisfying

$$n - d - x + y \leq r ,$$

that is,

$$y \leq x - n + d + r ,$$

and is represented in Figure 1 as the region on and below the line  $y = x - n + d + r$ . The intersection of the two spheres is then the region in the plane to the right of  $(x_0, y_0)$ , the point where the two boundary lines meet. Solving for  $x_0$  and  $y_0$  gives

$$x_0 = n - \frac{r+s+d}{2}$$

and

$$y_0 = \frac{r-s+d}{2} .$$

( $x_0$  and  $y_0$  may not be integers, but that does not matter.) We

can also write the equations of the two boundary lines as

$$y = y_0 - (x - x_0) \text{ and } y = y_0 + (x - x_0) .$$

Since the intersection of the spheres is the set of all points in S represented by the points in the right-hand region of the XY plane, the number of points in the intersection is

$$\sum_{x \geq x_0} \sum_{y=y_0-(x-x_0)}^{y=y_0+(x-x_0)} \binom{n-d}{x} \cdot \binom{d}{y} .$$

Since this double sum may be written as

$$\sum_x \binom{n-d}{x} \sum_y \binom{d}{y},$$

it is not as hard to evaluate numerically as other double sums would be. A computer algorithm for evaluating it could be written as follows: For each  $x \geq x_0$ , suppose that the inner sum (along a vertical line segment in the XY plane) has been computed. Multiply it by  $\binom{n-d}{x}$  and add it to the overall sum. Then increment  $x$ , update the inner sum over  $y$  by adding a term at each end of the sum, update  $\binom{n-d}{x}$ , and continue as above.

We will now transform the XY plane so that the joint probability distribution defined above for  $x$  and  $y$  may be approximated by the standard circular bivariate normal distribution. We define new variables by subtracting the mean and dividing by the standard deviation. Let

$$u = \frac{x - \frac{n-d}{2}}{\frac{\sqrt{n-d}}{2}}$$

and

$$v = \frac{y - \frac{d}{2}}{\frac{\sqrt{d}}{2}}.$$

Then  $u$  and  $v$  are independent random variables, each with mean 0 and standard deviation 1. Since for large  $n-d$  and  $d$ , each is approximately normal, their joint distribution is approximately the circular bivariate normal distribution, whose density function is

$$\frac{1}{2\pi} e^{-(u^2+v^2)/2} du dv.$$

The point  $(x_0, y_0)$  is transformed into  $(u_0, v_0)$  in the UV plane, as shown in Figure 2, where

$$u_0 = \frac{n-r-s}{\sqrt{n-d}}$$

and

$$v_0 = \frac{r-s}{\sqrt{d}}.$$

A little algebra shows that

$$u - u_0 = \frac{x - x_0}{\frac{\sqrt{n-d}}{2}}$$

and

$$v - v_0 = \frac{y - y_0}{\frac{\sqrt{d}}{2}}.$$

Thus the line  $y - y_0 = x - x_0$ , which borders the r-sphere and goes above the intersection, is transformed into

$$v - v_0 = \frac{\sqrt{n-d}}{\sqrt{d}} \cdot (u - u_0),$$

or  $v = v_0 + m(u - u_0)$ , a line in the UV plane whose slope is  $m = \frac{\sqrt{n-d}}{\sqrt{d}}$ . Similarly, the line bordering the s-sphere and going below the intersection becomes  $v = v_0 - m(u - u_0)$ . The region to the right of  $(u_0, v_0)$  in Figure 2 represents the intersection of the spheres.

We will now make some assumptions about the parameters; Figure 2 is drawn based on these assumptions. We assume that  $s \leq r < \frac{n}{2}$ , from which it follows that  $v_0 \geq 0$  and  $r+s < n$ , so that  $u_0 > 0$ . Thus  $(u_0, v_0)$  is in the first quadrant of the UV plane. Also, it follows that the line in Figure 2 bordering the r-sphere goes below the origin. (It must, because the volume of

the r-sphere is less than or equal to half of S.)

The volume of the intersection may be approximated by integrating the bivariate normal density over the region of the UV plane to the right of  $(u_0, v_0)$  and multiplying the result by  $2^n$ . There are many techniques for integrating this density over various regions in the plane. See for example Abramowitz and Stegun (1964), p. 956. We will derive an integral formula which can be transformed into Kanerva's formula if  $r = s$ .

We will describe points and lines in the UV plane using polar coordinates  $(R, \theta)$ , where  $u = R \cos \theta$  and  $v = R \sin \theta$ . The bivariate normal density function in polar coordinates becomes

$$\frac{1}{2\pi} e^{-R^2/2} R dR d\theta ,$$

which is circularly symmetric about the origin, since it is a function only of  $R$ .

Consider the line  $v = v_0 + m(u - u_0)$ , which is the boundary of the r-sphere. Draw a line segment from the origin to this line and perpendicular to it. Let  $A$  be the length of the segment, and let  $\alpha$  be the angle at the origin from the segment up to the U-axis (so that  $\alpha$  is a positive angle). The equation for the line bounding the r-sphere may now be written as

$$R = A \sec(\alpha + \theta) .$$

Since the angle between this line and the U-axis is  $\frac{\pi}{2} - \alpha$ , we have

$$m = \frac{\sqrt{n-d}}{\sqrt{d}} = \tan \left[ \frac{\pi}{2} - \alpha \right] = \cot \alpha = \frac{\cos \alpha}{\sin \alpha} ,$$

and since  $\sin^2 \alpha + \cos^2 \alpha = 1$ , we see that  $\cos \alpha = \frac{\sqrt{n-d}}{\sqrt{n}}$  and

$$\sin \alpha = \frac{\sqrt{d}}{\sqrt{n}}.$$

Let  $(R_0, \theta_0)$  be the point  $(u_0, v_0)$  expressed in polar coordinates. Since this point is on the line  $R = A \sec(\alpha + \theta)$ , we have

$$\begin{aligned} A &= R_0 \cos(\alpha + \theta_0) \\ &= R_0 \cos \theta_0 \cos \alpha - R_0 \sin \theta_0 \sin \alpha \\ &= u_0 \frac{\sqrt{n-d}}{\sqrt{n}} - v_0 \frac{\sqrt{d}}{\sqrt{n}} \\ &= \frac{(n-r-s) - (r-s)}{\sqrt{n}} \\ &= \frac{\frac{n}{2} - r}{\frac{\sqrt{n}}{2}}. \end{aligned}$$

Because of the circular symmetry of the density function, the integral of the density over the half-plane below the line  $R = A \sec(\alpha + \theta)$ , the region representing the  $r$ -sphere, is  $1-F(A)$ , or  $F(-A)$ , where  $F$  is the standard normal cumulative distribution function. To see this, rotate the plane so that the line bounding the  $r$ -sphere is vertical. It then crosses the new  $U$ -axis at  $A$ . Since the density function is circularly symmetric, the rotation does not change it, so we can go back to rectangular coordinates in the rotated plane and integrate over the region. Thus the volume of the  $r$ -sphere is approximately  $2^n \cdot F(-A)$ , as was shown by Kanerva (1984), p. 31.

Now draw a line from the origin through  $(u_0, v_0)$  and continue it beyond that point. Because of the assumptions we

made earlier, this line goes through the region representing the intersection of the spheres, cutting it into two pieces. We will evaluate the integral over each piece separately, beginning with the piece lying above the line we just drew.

We integrate over the upper piece as follows: Imagine a searchlight at the origin, rotating counterclockwise and sweeping over this piece of the plane. For each point  $(R, \theta)$  on the line  $R = A \sec(\alpha + \theta)$ , beginning with  $(R_0, \theta_0)$  and moving upward, we draw a ray from the origin through the point  $(R, \theta)$  and continuing through the region of integration. We integrate with respect to  $R$  along this ray (on which  $\theta$  is constant), as  $R$  increases from  $A \sec(\alpha + \theta)$  to  $\infty$ . Then we integrate with respect to  $\theta$ , as  $\theta$  increases from  $\theta_0$  to  $\frac{\pi}{2} - \alpha$ . The latter term is the upper limit for  $\theta$  because it represents a ray from the origin parallel to the line bounding the  $r$ -sphere. Thus the integral over the upper piece is

$$\int_{\theta=\theta_0}^{\frac{\pi}{2}-\alpha} d\theta \int_{R=A \sec(\alpha+\theta)}^{\infty} \frac{1}{2\pi} e^{-R^2/2} R \, dR .$$

The inner integral may be evaluated by letting  $t = R^2/2$ , giving us

$$\frac{1}{2\pi} e^{-A^2 \sec^2(\alpha+\theta)/2} .$$

Therefore, if we let  $z = \alpha + \theta$ , we can write the integral over the upper piece as

$$\frac{1}{2\pi} \int_{\alpha+\theta_0}^{\pi/2} e^{-A^2 \sec^2 z/2} dz .$$

To find the integral over the lower piece of the intersection, which lies between the line from the origin through  $(u_0, v_0)$  and the line bounding the s-sphere, we go through a similar process. Draw a line segment from the origin to the line bounding the s-sphere and perpendicular to it. The angle from the U-axis to this segment is  $\alpha$ , the same angle as before, but in the first quadrant. Let B be the length of the segment. The equation in polar coordinates of the line bounding the s-sphere is then  $R = B \sec (\theta - \alpha) = B \sec (\alpha - \theta)$ , and we can show that

$$B = \frac{\frac{n}{2} - s}{\frac{\sqrt{n}}{2}} .$$

Integrating with respect to R along a ray from the origin going through the lower piece of the intersection, and then integrating with respect to  $\theta$ , we have

$$\begin{aligned} \int_{\theta=\alpha-\frac{\pi}{2}}^{\theta_0} d\theta \int_{R=B \sec (\alpha-\theta)}^{\infty} \frac{1}{2\pi} e^{-R^2/2} R dR \\ = \frac{1}{2\pi} \int_{\alpha-\frac{\pi}{2}}^{\theta_0} e^{-B^2 \sec^2 (\alpha-\theta)/2} d\theta . \end{aligned}$$

The lower limit for  $\theta$  represents a ray from the origin parallel to the line bounding the s-sphere. Letting  $z = \alpha - \theta$  and interchanging the limits of integration, we have



$$\frac{1}{2\pi} \int_{\alpha - \theta_0}^{\pi/2} e^{-B^2 \sec^2 z/2} dz .$$

This integral is of the same form as the integral for the upper piece.

The approximation to the volume of the intersection of the spheres is then  $2^n$  times the sum of the two integrals.

In the special case  $r = s$ , we have  $v_0 = 0$ ,  $\theta_0 = 0$ , and  $A = B$ , so the line cutting the intersection into two pieces is the U-axis. The two pieces are therefore symmetric to each other, and their integrals are equal. The sum of the two integrals is therefore

$$\frac{1}{\pi} \int_{\alpha}^{\pi/2} e^{-A^2 \sec^2 z/2} dz .$$

Kanerva (1984), p. 157, derived the following integral approximation for this same quantity:

$$\int_{d/n}^1 \frac{1}{2\pi\sqrt{x(1-x)}} e^{-\frac{1}{2} \cdot \frac{c_p^2}{1-x}} dx ,$$

where

$$c_p = \frac{r - \frac{n}{2}}{\frac{\sqrt{n}}{2}} = -A .$$

We will show that these two integrals are equal.

Beginning with Kanerva's integral, let  $y = 1-2x$ . Then  $dy = -2dx$  and  $4x(1-x) = 1-(1-2x)^2 = 1-y^2$ . Now let  $y = \cos 2z$ . Then  $dy = -2 \sin 2z dz$ , and  $dx = \sin 2z dz$ . We note that the

function  $z = \frac{1}{2} \cos^{-1}(1-2x)$  is well-defined, continuous, and monotonically increasing for  $0 \leq x \leq 1$ , with range  $0 \leq z \leq \frac{\pi}{2}$ .

Also,

$$2\sqrt{x(1-x)} = \sqrt{1-y^2} = \sin 2z .$$

In the exponent we have the term

$$2(1-x) = 1 + y = 1 + \cos 2z = 2 \cos^2 z ,$$

where we used the double-angle formula for  $\cos 2z$ .

Kanerva's integral now becomes

$$\begin{aligned} \frac{1}{\pi} \int \frac{1}{\sin 2z} e^{-A^2/(2 \cos^2 z)} \sin 2z \, dz \\ = \frac{1}{\pi} \int e^{-A^2 \sec^2 z/2} \, dz . \end{aligned}$$

The upper limit of integration,  $x = 1$ , corresponds to  $y = -1$ , and therefore to  $z = \frac{\pi}{2}$ . The lower limit of integration is  $x = \frac{d}{n}$ . Since we found earlier that

$$\sin \alpha = \frac{\sqrt{d}}{\sqrt{n}} ,$$

the lower limit corresponds to

$$y = 1 - 2 \cdot \frac{d}{n} = 1 - 2 \sin^2 \alpha = \cos 2\alpha .$$

Finally, since  $y = \cos 2z$ , the lower limit is  $z = \alpha$ . Therefore the two integrals agree.

Kanerva (1984) states, on p. 158, that if the lower limit in his integral is 0, the integral is  $F(-A)$ , the approximation for the volume of the entire sphere. A lower limit of  $x = 0$  corresponds to  $z = 0$  in the equivalent integral above. Since we defined  $z$  to be  $\alpha + \theta$ , this integral sweeps over the part of the  $r$ -sphere in Figure 2 lying between  $\theta = -\alpha$  and  $\theta = \frac{\pi}{2} - \alpha$ .

Thus the region over which we are integrating is exactly one half of the half-plane which represents the  $r$ -sphere. The integral of the density function over this region is therefore half of  $F(-A)$ . Since our equivalent of Kanerva's integral is twice the integral of the density function, this integral equals  $F(-A)$ .

Kanerva (1984) then considers, on p. 159, the case where  $d$ , the distance between the centers, is  $\frac{n}{2}$ . In this case the spheres should be "independent", in the sense that the probability that a randomly chosen point lies in the intersection of the two spheres is approximately the product of the probabilities that the point lies in the individual spheres. (Of course, if the spheres are small, they would be disjoint, so we assume that the intersection is large enough so that these approximations may be applied.) We will consider the general case where  $r$  and  $s$  may be unequal. These probabilities may be approximated by integrating the density function over the appropriate regions. For the  $r$ -sphere, we saw that the integral is  $F(-A)$ . Similarly, for the  $s$ -sphere the integral is  $F(-B)$ .

Since  $d = \frac{n}{2}$ , we have  $m = 1$ , so the lines in the  $UV$  plane bounding the spheres have slopes of 1 and -1, and are therefore perpendicular. If we rotate the plane  $45^\circ$  counterclockwise, the line bounding the  $r$ -sphere will be vertical, crossing the new  $U$ -axis at  $A$ , and the line bounding the  $s$ -sphere will be horizontal, crossing the new  $V$ -axis at  $B$ . The intersection of the spheres is now represented by the part of the plane above and to the right of these lines. By the circular symmetry of the density function, the new  $u$  and  $v$  are independent normally

distributed random variables. Therefore, if we go back to rectangular coordinates and integrate over the intersection region, we find that the integral is  $F(-A) \cdot F(-B)$ , the product of the probabilities.

The integral approximations in this section may be improved upon by using the "continuity correction", a method often used when approximating discrete random variables by continuous ones. The correction consists of taking the radii of the  $r$ -sphere and the  $s$ -sphere to be  $r + \frac{1}{2}$  and  $s + \frac{1}{2}$ , respectively, so that the lines bounding the spheres in Figure 1 lie halfway between the points in the spheres and the points outside of the spheres. The equations of these lines, and the other quantities calculated from them, are then modified accordingly.

I wish to thank Mike Raugh for many stimulating discussions on this subject.

#### REFERENCES

- M. Abramowitz and I. E. Stegun (Eds.). 1965. *Handbook of Mathematical Functions*. Dover Publications.
- W. Feller. 1957. *An Introduction to Probability Theory and Its Applications, Volume 1* (2nd edition). John Wiley & Sons.
- P. Kanerva. 1984. *Self-Propagating Search: A Unified Theory of Memory*. Stanford University Ph.D. dissertation; Bradford Books (MIT Press). In press.

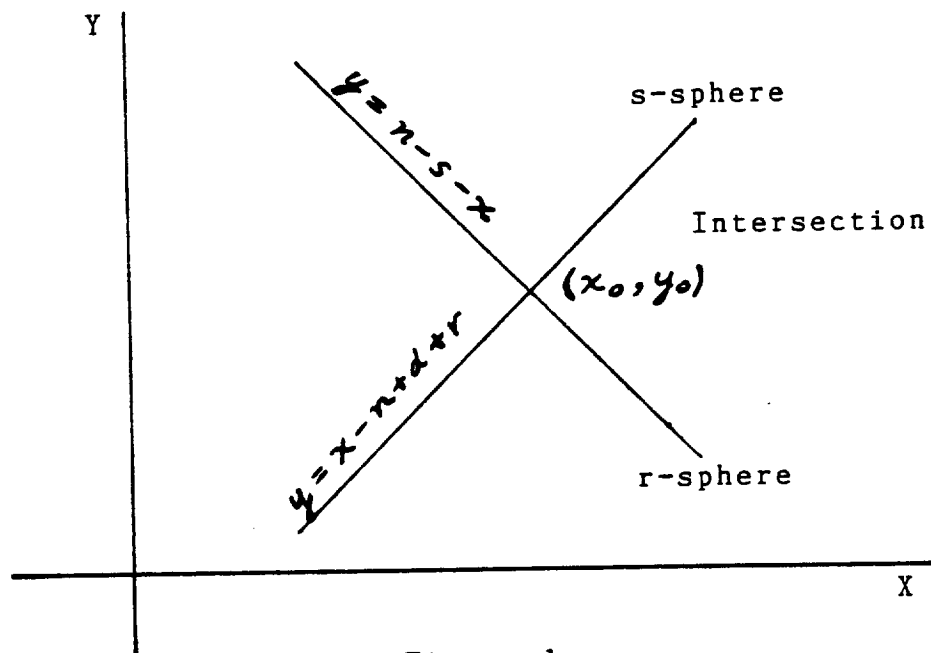


Figure 1

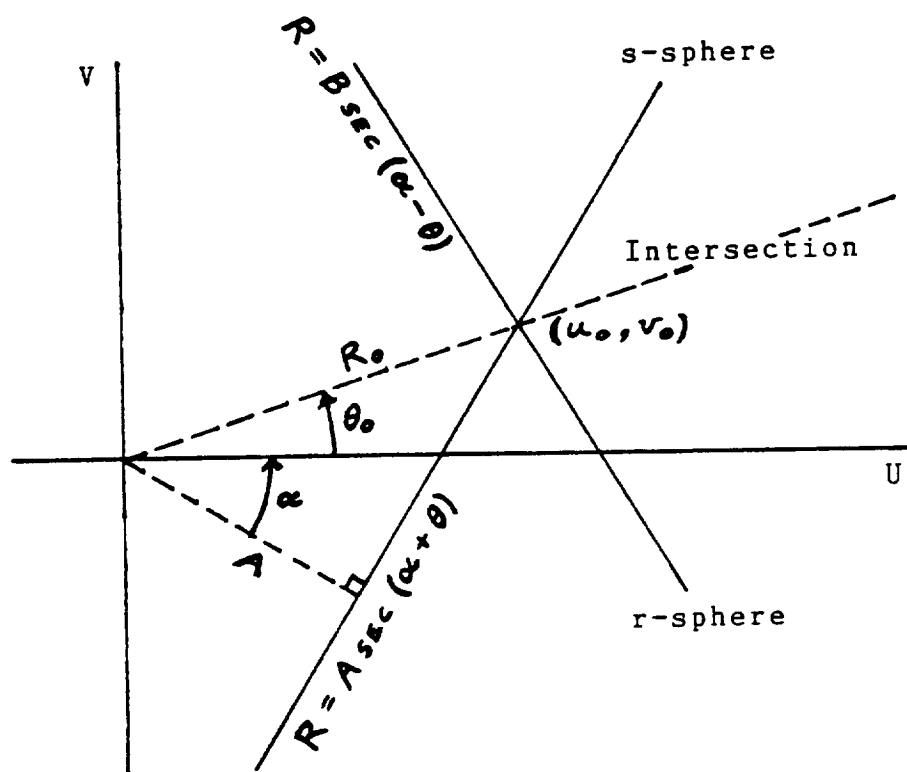


Figure 2

